Thresholds for Graph Properties and a Proof of the Kahn-Kalai Conjecture

Noémie Claret

London School of Economics and Political Science

March 28, 2024

Overview

Discrete Probability Space

Discrete Probability Space

A *discrete probability space* is a triple $(\Omega, \mathcal{A}, \mathbb{P})$ consisting of

1. the *sample space* Ω, which is the set of all possible outcomes of the random experiment,
2. an *algebra* $\mathcal{A}$ representing the allowable events, which is family of subset of Ω containing the empty set and closed under complement and finite union,
3. and a probability function $\mathbb{P} : \mathcal{A} \to [0, 1]$.

[Mitzenmacher, M. and E. Upfal (2005)].

Tail Bounds

Markov's Inequality

Let $Y \geq 0$ be a random variable on a discrete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\alpha > 0$. Then $\mathbb{P}(Y \geq \alpha) \leq \frac{\mathbb{E}(Y)}{\alpha}$. In particular, $\mathbb{P}(Y \geq 1) \leq \mathbb{E}(Y)$

Chebyshev's Inequality

Let Y be a random variable on a discrete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}(Y) < \infty$ and $Var(Y) < \infty$. Then for $\alpha > 0$,
$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq \alpha) \leq \frac{Var(Y)}{\alpha^2}$.

[Mitzenmacher, M. and E. Upfal (2005)].

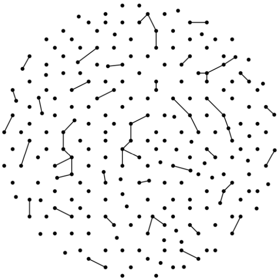


Figure: A Random Graph $G_{n,p}$ under the Erdős-Rényi Model. [Park, J. (2023)]

Consider the graph $G_{n,p}$ where each edge ω is drawn uniformly from the set of edges X with probability p. This is known as the Erdős-Rényi Model.

Random Graph on 3 vertices

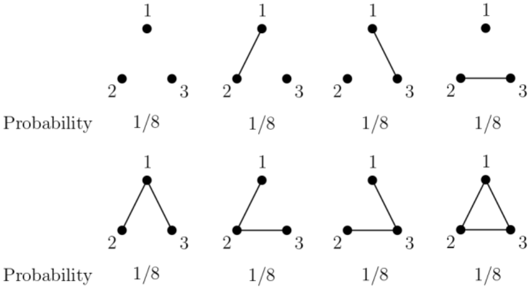


Figure: Probability Distribution of $G_{3,1/2}$. [Park, J. (2023)]

Each of the eight possible graph is is equally likely as each edge is drawn with probability $1/2$. However, the probability distribution of $G_{3,p}$ will look different as we change p.

Random Graph on 3 vertices -ctd

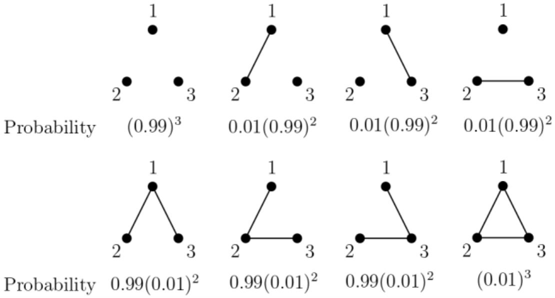


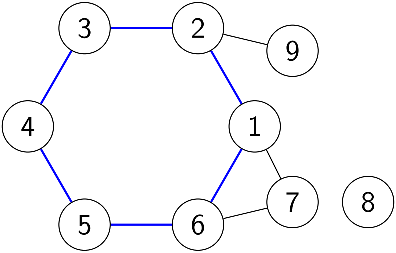
Figure: Probability Distribution of $G_{3,0.01}$ [Park, J. (2023)].

Natural question: What p should be chosen so that $G_{n,p}$ admits a property almost surely?

Properties in Random Graphs

Natural question: What p should be chosen so that $G_{n,p}$ admits a property almost surely?

- $G_{n,p}$ contains a triangle almost surely? $\rightarrow p \geq O(1/n)$ [Frieze, A. M. and Karoński (2015)]
- $G_{n,p}$ contains a Hamiltonian Cycle almost surely? $\rightarrow p \geq O(\log n/n)$ [Pósa, L. (1976)]



6-cycle

Figure: 6-cycle in $G(9, p)$

Example:

- Each cycle is a fixed set of k edges C_k in $G_{n,p}$. Denote $G_{n,p} \in \mathcal{F}$ the event 'for some k, the random graph $G_{n,p}$ contains a k-cycle on the potential set of edges C_k'.
- Define $Y_k^{(n)}$ to be the number of k-cycles cycles in the obtained graph, with $3 \leq k \leq n$.

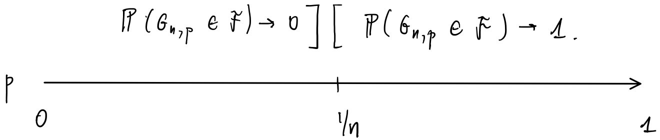


Figure: Location of Threshold $p^*(\mathcal{F})$.

$$\lim_{n\to\infty} \mathbb{P}(G_{n,p} \in \mathcal{F}) = \begin{cases} 0, & \text{if } p(n) \ll 1/n \\ 1, & \text{if } p(n) \gg 1/n \end{cases}$$

Notation:

- $p(n) \ll q(n)$ if $p(n)/q(n) \to 0$ as $n \to \infty$
- $p(n) \gg q(n)$ if $p(n)/q(n) \to \infty$ as $n \to \infty$

k-Cycles in Random-Graphs

- $Y^{(n)}$ the number of all possible sets of k edges $Y_k^{(n)} = \sum_{C_k} 1_{C_k \in \mathcal{F}}$ is obtained by marginalizing over k.

$$Y^n = \sum_{k=3} Y_k^{(n)} = \sum_{k=3}^{n} \sum_{C_k} 1_{C_k \in \mathcal{F}}$$

- Denote $(n)_k := n(n-1)...(n-k+1)$.
- There are $\frac{(n)_k}{2k} = \binom{n}{k}\frac{(k-1)!}{2}$ possible k-cycles.
- For a given $G_{n,p}$, $1_{C_k \in \mathcal{F}}$ is Bernoulli with parameter p^k.

$$\mathbb{E}(Y^{(n)}) = \sum_{k=3}^{n} \mathbb{E}(Y_k^{(n)}) = \sum_{k=3}^{n} \frac{(n)_k}{2k} p^k < (np)^3 \sum_{k=0}^{n-3} (np)^k = \frac{(np)^n - (np)}{np - 1}$$

By linearity of expectation, and noting $\binom{n}{k}(k-1)! < \binom{n}{k}k! < n^k$.

'First Moment Method': Markov's Inequality:

$$\lim_{n\to\infty} \mathbb{P}(Y^{(n)} \geq 1) \leq \lim_{n\to\infty} \mathbb{E}(Y^{(n)}) \leq \lim_{n\to\infty} \frac{(np)^n - (np)^3}{np - 1}$$

$$1 - \lim_{n\to\infty} \mathbb{P}(Y^{(n)} \geq 1) \geq 1 - \lim_{n\to\infty} \mathbb{E}(Y^{(n)}) \geq 1 - \lim_{n\to\infty} \frac{(np)^n - (np)^3}{np - 1}$$

- Using Markov's inequality, we can show that $G_{n,p} \notin \mathcal{F}$ if p is such that $\lim_{n\to\infty} np = 0$.
- Hence if p is such that $\lim_{n\to\infty} \frac{(np)^n}{1-np} = 0$, then $\lim_{n\to\infty} \mathbb{P}(G_{n,p} \in \mathcal{F}) = 0$, which is the case if $p \leq (1/n)$.

'Second Moment Method': Chebyshev's Inequality: We would like to show that for $p \gg 1/n$ $G_{n,p} \in \mathcal{F}$. In particular, let $p = \frac{10}{n}$.

- A sufficient condition for $G_{n,p} \in \mathcal{F}$ is that $G_{n,p}$ contains n edges.
- The number of edges $Z^{(n)} = \sum_{i=1}^{n} E_i$ is distributed as $bin(\binom{n}{2}, p)$.
- $\mathbb{E}(Z^{(n)}) = \binom{n}{2}\frac{10}{n} = \frac{10}{n}\left(1 - \frac{1}{n}\right)$, $Var(Z^{(n)}) = \binom{n}{2}\frac{10}{n}$.

Using Chebyshev's inequality with $\alpha = \frac{10}{n}$,

$$\mathbb{P}(|\mathbb{E}(Z^{(n)}) - Z^{(n)}| \geq \frac{n}{10}) \leq \frac{Var(Z^{(n)})}{(n/10)^2} = \frac{10^3}{2n} \cdot (1 - \frac{1}{n}) \cdot (1 - \frac{10}{n}) \to 0.$$

$Z^{(n)} > \mathbb{E}(Z^{(n)}) - \frac{n}{10}$ a.s. implies $Z_n > \frac{49}{10}n - 5$, and it can easily be shown by induction that with $n \geq 3$, we have $Z^{(n)} \geq n$.

Monotone Increasing Property

Increasing Property

An *increasing property* $\mathcal{F} \subseteq 2^X$ is a family of sets such that if $B \supseteq A \in \mathcal{F}$, then $B \in \mathcal{F}$. A property is *non-trivial* if $\mathcal{F} \neq \emptyset$.

The event $X_p \in \mathcal{F}$

X_p is a uniformly drawn subset of X, where each element is taken with probability p. Denote by $X_p \in \mathcal{F}$ the event ' $\exists A \in \mathcal{F}$ s.t. $X_p \subseteq A$'.

$$\mathbb{P}(X_p \in \mathcal{F}) = \sum_{A \in \mathcal{F}} \mathbb{P}(A) = \sum_{A \in \mathcal{F}} p^{|A|}(1-p)^{|X \setminus A|}.$$

Threshold Functions

Threshold

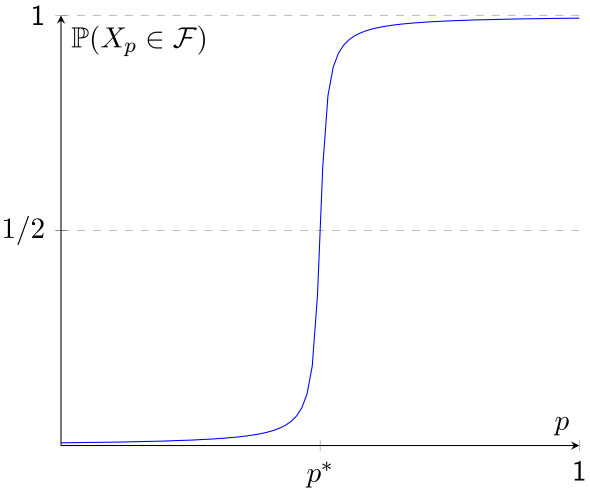
A function $p^* = p^*(n)$ is a *threshold* for a monotone increasing property $\mathcal{F} \subseteq 2^X$ is the unique p^* such that $\mathbb{P}(X_{p^*} \in \mathcal{F}) = 1/2$ a.s. .

Bollobás-Thomason Theorem

Let $\mathcal{F}$ be non-trivial increasing. For every $0 < \epsilon \leq 1/2$, there exists a unique p^* and some constants $K_\epsilon^1, K_\epsilon^2$ such that

$$\mathbb{P}(X_{p^*} \in \mathcal{F}) = 1/2,\ \mathbb{P}(X_{k_\epsilon^2 p^*} \notin \mathcal{F}) \leq \epsilon \text{ and } \mathbb{P}(X_{K_\epsilon^2 p^*} \in \mathcal{F}) \geq 1 - \epsilon.$$

Threshold Functions-ctd



Hamiltonian Cycles

Threshold for Hamiltonian Cycle is non-trivial.

$$\lim_{n\to\infty} \mathbb{P}(X_p \in \mathcal{F}) \begin{cases} = 0, & \text{if } p^*(n) \gg p(n) \\ \neq 1, & \text{if } p(n) \gg p^*(n) \end{cases}$$

- Y_n the number of Hamiltonian Cycles in $G_{n,p}$
- $\mathbb{E}(Y_n) = p^n(n-1)/2$
- Using Markov's Inequality,

$$\mathbb{P}(G_{n,p} \notin \mathcal{F}) = 1 - \lim_{n\to\infty} \mathbb{P}(Y_n \geq 1) \geq 1 - \lim_{n\to\infty} p^n(n-1)/2$$

- Markov's inequality implies $G_{n,p} \notin \mathcal{F}$ if $p \ll (n^{-1/n})$.
- However, it is not the case that $G_{n,p} \in \mathcal{F}$ if $p \gg (n^{-1/n})$!

[Pósa, L. (1976)] proved that for $\mathcal{F}$ given by '$G_{n,p}$' admits a Hamiltonian cycle, $p^*(\mathcal{F}) = \frac{\log n}{n}$.

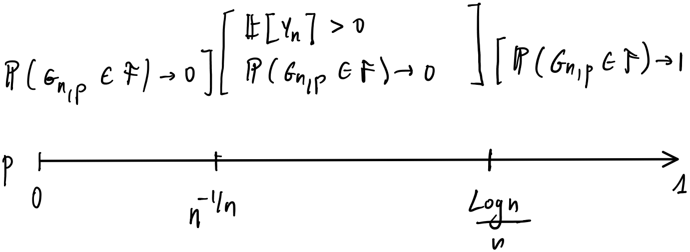


Figure: Location of $q(\mathcal{F})$ and $p^*(\mathcal{F})$.

$$\lim_{n\to\infty} \mathbb{P}(G_{n,p} \in \mathcal{F}) = \begin{cases} 0, & \text{if } p(n) \ll \frac{\log n}{n} \\ 1, & \text{if } p(n) \gg \frac{\log n}{n} \end{cases}$$

Finding Thresholds

- Given $\mathcal{F}$, we are interested in finding $p^*(\mathcal{F})$. This is generally a difficult task. While it ensures existence of p^*, the Bollobás-Thomason Theorem does not provide information on its location.
- One intuitive starting point to find these threshold function is the 'Expectation Threshold'.
- A key concept behind the expectation threshold is that of a p-*small cover* .

Minimal Element

A *minimal* element G of a set $\mathcal{G}$ is such that $\forall G' \in \mathcal{G}$, then $G \subsetneq G'$.

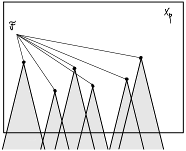


Figure: Increasing Property $\mathcal{F}$ and its Minimal Elements

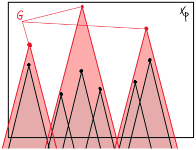


Figure: Minimal elements for a Cover $\mathcal{G}$ of $\mathcal{F}$

- The black dots represent the sets in $\mathcal{F}_{min} := F \in \mathcal{F} : F$ minimal.
- The points contained inside the cones is $\mathcal{F} = \langle \mathcal{F}_{min} \rangle$.
- $\mathcal{F} \subseteq \langle G \rangle$, but, crucially, the sets in G have less overlap.

p-Small Cover

p-Small Cover

We say that the property $\mathcal{F}$ is *p-small* if there exists $\mathcal{G} \subseteq 2^X$ such that both of the following hold.

$$\mathcal{F} \overset{(1)}{\subseteq} \langle \mathcal{G} \rangle := \bigcup_{S \in \mathcal{G}} \left\{ T \subseteq 2^X : S \subseteq T \right\} \text{ and } \sum_{S \in \mathcal{G}} p^{|S|} \overset{(2)}{\leq} \frac{1}{2}$$

If (1) holds, we say $\mathcal{G}$ is a *cover* of $\mathcal{F}$.

Remark

$\sum_{S \in \mathcal{G}} p^{|S|} = \mathbb{E}[|\{S \in \mathcal{G} : S \subseteq X_p\}|]$

- We are interested in the size of $\mathcal{F} := F \in \mathcal{F} : F \subseteq X_p$. The idea is to approximate the set $\mathcal{F}$ with one of lower expected 'size'.

Expectation Threshold

Expectation Threshold

The *expectation threshold* $q(\mathcal{F})$ of $\mathcal{F}$ is the maximum p such that $\mathcal{F}$ is p-small.

- One immediately notices that the expectation threshold provides a trivial lower bound on the threshold, as shown below.

$$\mathbb{P}(X_p \in \mathcal{F}) \leq \mathbb{P}(X_p \in \langle G \rangle) \leq \sum_{S \in \mathcal{G}} p^{|S|},$$

- For an arbitrary $p < q(\mathcal{F})$, $\mathbb{P}(X_p \in \mathcal{F}) < \frac{1}{2}$, preventing p from being the threshold p^*.

Location of p^*

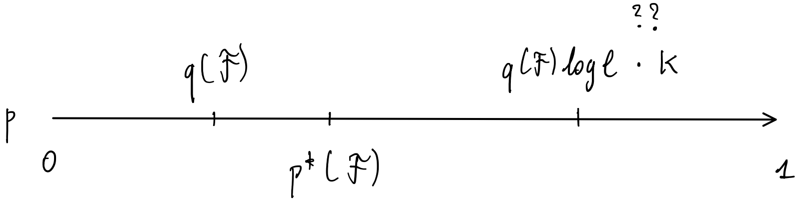


Figure: Location of $q(\mathcal{F})$ and $p^*(\mathcal{F})$.

[Kahn, J. and G. Kalai (2007)] conjectured that $p^*(\mathcal{F})$ is in fact bounded above by $q(\mathcal{F})$ multilied by some small constant, and stated that it would 'be more sensible to conjecture that it is not true'.

The Kahn-Kalai Conjecture

Kahn-Kalai Conjecture

For every finite set X and non-trivial increasing property $\mathcal{F} \subseteq 2^X$ with largest element of size l, there exists a universal constant K such that

$$p^*(\mathcal{F}) \leq Kq(\mathcal{F}) \log l(\mathcal{F}).$$

.[Kahn, J. and G. Kalai (2007)]

Reformulation

Let $l \geq 2$. There exists a universal constant L such that for any nonempty, l-bounded hypergraph $\mathcal{H}$ on X which is not p-small, a uniformly random $(Lp \log l)|X|$-element subset of X belongs to $\langle \mathcal{H} \rangle$ with probability $1 - o(1)$. [Park, J. and H. T. Pham (2022)]

The Kahn-Kalai Conjecture

Kahn-Kalai Conjecture

For every finite set X and non-trivial increasing property $\mathcal{F} \subseteq 2^X$ with largest element of size l, there exists a universal constant K such that

$$p^*(\mathcal{F}) \leq Kq(\mathcal{F}) \log l(\mathcal{F}).$$

.[Kahn, J. and G. Kalai (2007)]

Remark: The upper bound provided by the Kahn-Kalai Conjecture is in fact the threshold for Hamiltonian Cycles!

$$\lim_{n\to\infty} \mathbb{P}(G_{n,p} \in \mathcal{F}) = \begin{cases} 0, & \text{if } p(n) < \frac{\log n}{n} - \epsilon \\ 1, & \text{if } p(n) > \frac{\log n}{n} + \epsilon \end{cases}$$

Proof: Strategy

Strategy: iteratively construct two sets of elements: $\bigcup_i \mathcal{U}_i$ and $W = \bigcup_i W_i$ through a randomized iterative process. This algorithmic construction has two outputs.

1. $\bigcup_i \mathcal{U}_i$ is a p-small cover of $\mathcal{H}$.
2. W satisfied the conclusion of the reformulated conjecture: it is a uniformly random $(Lp \log l)|X|$-element subset of X belongs to $\langle \mathcal{H} \rangle$ w.h.p.

Proof: Minimum Fragments

- Say we have $W \in \binom{X}{w}$, and some hypergraph $\mathcal{H} \subseteq 2^X$.
- Given any element S of $\mathcal{H}$, then $T(S, W)$ is a *minimum (S, W)-fragment if $T(S, W) = S' \setminus W$ for some $S' \in \mathcal{H}$ such that $S' \subseteq S \cup W$ and $|S' \setminus W|$ is smallest.*
- *We denote the size of a* minimum fragment$|T(S, W)|$ *by $t(S, W)$.*

Algorithm 1: ConstructCover

Input : l-bounded hypergraph $\mathcal{H}$ which is not p-small, finite set X
Output: A family $\mathcal{U}$, a $m = (L_i p \log l)|X|$-sized random $W \in \binom{X}{m}$ subset of X.

$i := 0$, $X_0 := X$, $\mathcal{H}_0 := \mathcal{H}$, $\mathcal{U}_0 := \emptyset$, $\mathcal{G}_0 := \emptyset$;
while $\mathcal{H}_i \neq \{\emptyset\}$ *or* $\mathcal{H}_i \neq \emptyset$ **do**
 $W_i =$ uniformly drawn from $\binom{X_i}{m}$;
 $X_{i+1} \leftarrow X_i \setminus W_i$;
 $\mathcal{G}_i \leftarrow \{S_{i-1} \in \mathcal{H}_{i-1} : t(S_{i-1}, W_i) \geq 0.9 l_i\}$;
 $\mathcal{U}_i \leftarrow \{T(S_{i-1}, W_i) : S_{i-1} \in \mathcal{G}_i\}$;
 $\mathcal{H}_i \leftarrow \{T(S_{i-1}, W_i) : S_{i-1} \in \mathcal{H}_{i-1} \setminus \mathcal{G}_i\}$;
end
return $\bigcup_i \mathcal{U}_i$, $W = \bigcup_i W_i$

Squares of Hamiltonian Cycles

Power of Hamilton Cycle

The k^{th} power C^k of C is the cycle graph with vertex set $[n]$ and edge set $E(H) = uv : d_H(u, v) \leq k$, where $d_H(u, v) \leq k$ is the length of a shortest u, v-path in C if one exists. [Montgomery, R. (2018)]

- [Riordan, O. (2000)] showed that for $k \geq 3$, if $p = O(n^{-1/k})$, then $G_{n,p}$ contains a k-cycle C_n^k a.s. using an argument based on the 'Second Moment' method.
- [Kühn, D. and D., Osthus (2012)] conjectured that the threshold is $n^{-1/2}$
- The best bounds before 2020 were of $(\log n)^3 n^{-1/2}$ by [Fischer, M. and N. and S. Škorić (2022)] and of $(\log n)^2 n^{-1/2}$ by [Friedgut, E. (2005)].

Squares of Hamiltonian Cycles

- Let $\mathcal{H}_{C_n^2}$ be the hypergraph of all graphs on n vertices containing C^2.
- [Frieze, A. M. and Karoński (2015)] show that $\mathcal{H}_{C_n^2}$ is not $(e^{-1}n^{1/2})^{-1}$ – small.
- Noting that the largest element of $\mathcal{H}_{C_n^2}$ has size at most $l(\mathcal{H}_{C_n^2}) := n + \lfloor n/2 \rfloor = O(n)$.
- The Kahn-Kalai Conjecture (Theorem) implies there exists a universal constant K such that

$$n^{-1/2} \leq Kn^{-1/2} \log l(\mathcal{H}_{C_n^2}) = O(n^{-1/2} \log n).$$

References

Mitzenmacher, M. and E. Upfal (2005)
Randomized algorithms and probabilistic analysis
Cambridge University Press, Cambridge xvi+352.

Riordan, O. (2000)
Spanning subgraphs of random graphs
Combinatorics, Probability and Computing 125–148.

Kühn, D. and D., Osthus (2012)
On Pósa's conjecture for random graphs
SIAM Journal on Discrete Mathematics 1440–1457.

KFischer, M. and N. and S. Škorić(2022)
Triangle resilience of the square of a Hamilton cycle in random graphs
Journal of Combinatorial Theory. Series B 171–220.

References

KF̈riedgut, E. (2005)
Hunting for sharp thresholds
JRandom Structures & Algorithms 37–51.

Montgomery, R. (2018)
Topics in Random Graphs

Frieze, A. and Karoński, M. (2015)
Introduction to Random Graphs
Cambridge University Press.

Kahn, J. and Kalai, G. (2007)
Thresholds and expectation thresholds
Combinatorics, Probability and Computing.

References

Park, J. and Pham, H. T. (2022)
A proof of the Kahn-Kalai conjecture
2022 IEEE 63rd Annual Symposium on Foundations of Computer Science—FOCS 2022 636–639.

Park, J. (2023)
Threshold phenomena for random discrete structurese
Notices of the American Mathematical Society.

Pósa, L. (1976)
Hamiltonian circuits in random graphs
Discrete Mathematics 359–364.

Anup Rao (2023)
The Sunflower Lemma and Monotone Thresholds
https://www.youtube.com/watch?v=Rw6FEZnuE5Y.